Exploratory Data Analysis Report – Hotel Booking Dataset

Univariate Analysis

- Distribution of key metrics like ADR, lead time, booking changes.
- Visualized using histograms and density plots.

Bivariate Analysis

- Compared ADR across hotel types, customer segments, and market channels.
- Used boxplots and bar charts for comparisons.

Multivariate Analysis

- Correlation analysis using heatmaps.
- Pairwise relationships explored using pair plots.

Time Series Analysis

- Monthly trends in bookings and cancellations.
- Seasonal patterns in ADR and room demand.

Hypothesis Testing

- T-tests and ANOVA to validate ADR differences across categories.
- Chi-square test for associations between country and booking status.
- Z-tests for proportion-based comparisons.

Key Business Questions Addressed

1. **What influences ADR the most?**

- Lead time, hotel type, and customer category.

2. **Do guests who book earlier tend to request more changes?**

   - Yes, supported by correlation and visualization.

3. **Are there pricing or booking differences across countries?**

   - Yes, ADR and booking frequency vary significantly by country.

4. **Is there a pattern in room upgrades or mismatches?**

   - Mostly observed during peak seasons or high lead-time bookings.

Insights & Recommendations

- Leverage high-ADR customer segments with targeted promotions.

- Automate early booking change handling to improve service efficiency.

- Adapt pricing dynamically based on country-specific trends.

- Reduce room mismatches to boost guest satisfaction and reviews.

Conclusion

This exploratory data analysis provided deep insights into hotel booking behaviors, guest profiles, and revenue-impacting variables. By identifying patterns in ADR, booking lead time, cancellations, and country-specific trends, the analysis enables data-driven decision-making to optimize revenue, enhance customer satisfaction, and streamline operations. Further predictive modeling could be the next step to forecast cancellations or high-ADR bookings.